

Optimal risk sharing under distorted probabilities

Michael Ludkovski · Virginia R. Young

Received: 19 September 2008 / Accepted: 25 May 2009 / Published online: 16 June 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract We study optimal risk sharing among n agents endowed with distortion risk measures. Our model includes market frictions that can either represent linear transaction costs or risk premia charged by a clearing house for the agents. Risk sharing under third-party constraints is also considered. We obtain an explicit formula for Pareto optimal allocations. In particular, we find that a stop-loss or deductible risk sharing is optimal in the case of two agents and several common distortion functions. This extends recent result of Jouini et al. (Adv Math Econ 9:49–72, 2006) to the problem with unbounded risks and market frictions.

Keywords Distortion risk measures · Comonotonicity · Risk sharing · Pareto optimal allocations

Mathematics Subject Classification (2000) 91B30 · 91B32 · 62P05

1 Introduction

Many financial problems involve transfer of risk among agents. Two noteworthy examples are insurance markets and the general equilibrium theory of stock prices. In such problems, $n \geq 2$ agents with risky endowments (or loss exposures) X_i for $i = 1, 2, \dots, n$ are interested in devising an optimal re-allocation of their risks. Let $X \triangleq \sum_{i=1}^n X_i$ be the total exposure of the n agents, and let V_i be the subjective valuation (preference) functional of the i th agent.

M. Ludkovski (✉)
Department of Statistics and Applied Probability,
University of California, Santa Barbara, CA 93106, USA
e-mail: ludkovski@pstat.ucsb.edu

V. R. Young
Department of Mathematics, University of Michigan,
530 Church St., Ann Arbor, MI 48109, USA
e-mail: vryoung@umich.edu

Consider the collection of allocations of the loss X , namely

$$\mathcal{A}(X) \triangleq \left\{ \mathbf{Y} := (Y_1, Y_2, \dots, Y_n) : X = \sum_{i=1}^n Y_i, \ V_i(Y_i) \text{ finite} \right\}.$$

The risk sharing problem consists in finding an *optimal* allocation $\mathbf{Y}^* \in \mathcal{A}(X)$, namely an allocation such that (i) \mathbf{Y}^* is *Pareto optimal*, that is, no agent can be made strictly better off without another agent being made strictly worse off; and (ii) \mathbf{Y}^* satisfies a *rationality constraint*, that is, all agents are at least as well off under \mathbf{Y}^* as under the initial exposures $\mathbf{X} = (X_1, X_2, \dots, X_n)$. The latter feasibility constraint is motivated by the assumption that only an irrational agent would enter into a contract that made the agent (strictly) worse off.

The key ingredient in the above problem are the preference functionals V_i , and accordingly the optimal risk sharing literature has evolved as new theories of risk have been developed. Pioneering work was carried out in the 1960s by Borch [7] and Arrow [3] who showed that deductible insurance is optimal under concave risk preferences, specifically, when V_i are represented by von Neumann–Morgenstern utility functions. Borch showed that the optimal risk sharing arrangement was solely a function of the total risk X , a type of mutual fund result. More generally, Wilson [32] showed that when agents are expected utility maximizers, individual risks do not matter, and agents optimally consume a non-decreasing function of the total risk X . We also obtain a mutual fund principle within the framework of our model.

Later research studied the case of the dual theory of risk of Yaari (Young and Browne [34]), Choquet expected utility theory (e.g. Chateauneuf et al. [11]) and rank dependent utility (see Carlier and Dana [9, 10]). Very recently, research has focused on risk preferences given in terms of convex risk measures, as exemplified by the monograph of Föllmer and Schied [17]. In particular, Barrieu and El Karoui [4] studied optimal risk sharing under the exponential indifference measure, while Jouini et al. [20] analyzed the case of two agents and convex, law-invariant risk measures. The related question of market equilibrium was addressed by Acciaio [2], Burgert and Rüschendorf [8] and Filipovic and Kupper [16]. On a more abstract level, Ludkovski and Rüschendorf [22] show that Pareto optimal allocations are comonotone if the risk measures preserve the convex order; also see the earlier work of Landsberger and Meilijson [21] and Dana and Meilijson [12]. The latter structural result allows for some explicit computations, as it permits direct representation of possible allocations through the pooling functions.

A simultaneous strand of the literature has been addressing extensions of the basic insurance problem that take into account market frictions. For example, the fundamental problem of adverse selection was initiated by Rothschild and Stiglitz [28] and later further discussed in Young and Browne [34]. The effect of transaction costs on optimal contracts was first considered by Raviv [25]. Other possible externalities are summarized in the survey articles of Gerber [18] and Aase [1]. Many markets also impose constraints on possible risk transfers. Often, only a limited set of risk instruments is a priori given, so that risk sharing must belong to the span of available contracts (as studied by Filipovic and Kupper [15]). Alternatively, the amount of risk transfer is limited by regulator authorities; for instance in the classical insurance problem the insurer may be able to take on only part of the total risk due to risk capital regulations. The latter problem, which we call risk sharing under constraints, introduces effectively $n + 1$ players into the model, namely n original participants, plus the additional regulator that imposes limits on allowable risk exposures of each participant. The special case of Value-at-Risk constraints was recently analyzed in Bernard and Tian [6].

This article extends previous results in these two directions by studying optimal risk sharing in the context of distortion risk measures, transaction costs and/or third-party constraints.

Distortion risk measures lie at the junction of actuarial and financial applications, being related both to the dual theory of risk and coherent risk measures. The transaction costs in our model have a dual nature and can either represent genuine transaction fees arising due to verification, accounting and other inter-agent costs, or the risk-loaded premium charged by the insurer. For the constraints, we consider a general set of restrictions given in terms of distortion risk measures.

Our main result, namely Theorem 2, shows that in all of the above cases, the optimal risk allocation consists of a collection or “ladder” of deductible contracts. This result can be interpreted as an economic justification for the tranche (or call-spread) contracts one observes in practice, in particular, in credit and reinsurance markets. Moreover, using the quantile representation of distortion risk measures we are able to explicitly characterize Pareto optimal contracts under transaction costs and/or constraints. In turn, this allows us to present several completely worked-out examples of optimal risk sharing under some common risk measures, such as Average Value-at-Risk.

In terms of related literature, Theorem 2 is an extension of the results of Jouini et al. [20] to the multi-agent case with transaction costs and constraints. Compared to their abstract approach based on convex duality an inf-convolution, our method is more elementary and direct and provides a clearer insight into the problem structure. On a more general note, this paper aims to underscore the usefulness of distortion risk measures that have been arguably under-appreciated by the financial/mathematical economics community, as pointed out by Denuit et al. [13]. In contrast to the classical expected utility theory, this new framework is driven by two factors. First, it postulates cash-equivariant preferences that are appealing based on the normative observation that guaranteed cash payments should not affect risk attitudes. Secondly, distortion risk measures attempt to mirror business practices where various Value-at-Risk (VaR) methodologies have emerged as the tool of choice. In particular, Average Value-at-Risk (AVaR) has been gaining practitioner acceptance and also happens to be a canonical example of our model.

This paper is organized as follows: In Sect. 2, we define the setting in which the n agents seek a Pareto optimal risk exchange. In Sect. 3, we obtain the class of Pareto optimal risk exchanges in our model. This is then generalized to the constrained setting in Sect. 4. We focus on the case of two agents in Sect. 5, while interpreting one agent as an insurer and another as a buyer of insurance. In this simplified setup we present fully solved examples, including examples with explicitly computable deductibles. In Sect. 6, we provide another illustration of our results by considering a single-agent minimization by a buyer of insurance who faces a regulator constraint on the possible indemnity contracts. Section 7 concludes the paper.

2 Model for risk sharing

2.1 Distorted probabilities

Consider the collection of a.s.-finite random variables $\mathcal{P} = \{Y : \mathbb{P}[-\infty < Y < \infty] = 1\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. As usual, we denote by $L^\infty \subset \mathcal{P}$ ($L^1 \subset \mathcal{P}$) the collection of all a.s. bounded (respectively integrable) random variables.

Definition 1 Two random variables Y and $Z \in \mathcal{P}$ are said to be comonotone if

$$(Y(\omega_1) - Y(\omega_2))(Z(\omega_1) - Z(\omega_2)) \geq 0, \quad (1)$$

$\mathbb{P}(d\omega_1) \times \mathbb{P}(d\omega_2)$ -almost surely. In other words, Y and Z move together.

An equivalent definition of comonotonicity is that there exists a random variable $V \in \mathcal{P}$ and non-decreasing functions f_Y and f_Z such that $Y = f_Y(V)$ and $Z = f_Z(V)$ almost surely (see Denneberg [14]). Another equivalent definition is that there exist non-decreasing functions h_Y and h_Z such that $h_Y(x) + h_Z(x) = x$, $Y = h_Y(Y + Z)$, and $Z = h_Z(Y + Z)$ almost surely.

Denote by S_Y the (decumulative) distribution function of Y , that is, $S_Y(t) = \mathbb{P}(Y > t)$, and by S_Y^{-1} the (pseudo-)inverse of S_Y , which is unique up to a countable set. For concreteness, take $S_Y^{-1}(p) = \sup\{t : S_Y(t) > p\}$. The inverse S_Y^{-1} thus defined is right continuous; if one were to desire left continuity, then replace $>$ with \geq .

Definition 2 Let $g : [0, 1] \rightarrow [0, 1]$ be a non-decreasing, concave function such that $g(0) = 0$, $g(1) = 1$. The distortion risk measure H_g is defined as

$$\begin{aligned} H_g(Y) &= \int Y d(g \circ \mathbb{P}) = \int_0^1 S_Y^{-1}(p) dg(p) \\ &= \int_{-\infty}^0 (g[S_Y(t)] - 1) dt + \int_0^{\infty} g[S_Y(t)] dt, \quad \forall Y \in \mathcal{P}. \end{aligned} \quad (2)$$

The function g is called a *distortion* because it modifies, or distorts, the tail probability S_Y . Observe if $g(p) = p$, then $H_g(Y) = \mathbb{E}Y$. For this reason, H_g is also referred to as an expectation with respect to a distorted probability. Note that at this stage we allow H_g to take $\pm\infty$ as a value.

We assume that each agent orders random variables in \mathcal{P} by using a distortion risk measure H_g , where Y is preferred to (that is, less risky than) Z by the agent if $H_g(Y) \leq H_g(Z)$, and we pursue this topic in the next section. For more background on such risk measures H_g , see Yaari [33] who discusses evaluating random variables in a theory of risk that is dual to expected utility. See also Wang et al. [31] for an axiomatic definition of distortion risk measures as law-invariant, coherent and comonotone-additive functionals on \mathcal{P} . Two noteworthy examples of distortion risk measures are (1) the Average Value-at-Risk at level $1 - \alpha^{-1}$ (AVaR) obtained by taking $g(p) = \min(\alpha p, 1)$ for some $\alpha > 1$ and (2) the proportional hazards transform $g(p) = p^c$ for some $0 < c < 1$.

Definition 3 Y is said to precede (or be preferred to) Z in *convex order* if $\int_0^q S_Y^{-1}(p) dp \leq \int_0^q S_Z^{-1}(p) dp$ for all $q \in [0, 1]$ with equality at $q = 1$. We write $Y \leq_{cx} Z$.

Note that convex order is equivalent to ordering with respect to second stochastic dominance with equal means, as originally shown in Rothschild and Stiglitz [26, 27]. For later use, recall that H_g satisfies the following properties for $Y \in \mathcal{P}$ (see Wang and Young [30]):

- (a) $H(Y)$ depends only on the law of $Y \in \mathcal{P}$.
- (b) H is cash equivariant: $H(Y + a) = H(Y) + a$ for any $a \in \mathbb{R}$.
- (c) H is subadditive in general and additive for comonotone risks:

$$H(Y + Z) \leq H(Y) + H(Z), \quad \text{for } Y, Z \in \mathcal{P}, \quad (3)$$

with equality for any Y, Z comonotone.

- (d) Positive homogeneity: If $a \geq 0$, then $H_g(aY) = aH_g(Y)$. Together with (c) this implies that H is convex, that is, $H(\lambda Y + (1 - \lambda)Z) \leq \lambda H(Y) + (1 - \lambda)H(Z)$ for all $\lambda \in (0, 1)$.

- (e) Duality: $H_g(-Y) = -H_{\tilde{g}}(Y)$, in which \tilde{g} is the dual distortion of g given by $\tilde{g}(p) = 1 - g(1 - p)$ for $p \in [0, 1]$. The dual distortion \tilde{g} is convex and $H_{\tilde{g}}$ can be thought of as a monetary utility function, see Jouini et al. [19].
- (f) Convex ordering: H_g preserves \leq_{cx} . If $Y \leq_{cx} Z$ then $H_g(Y) \leq H_g(Z)$. In particular, because $\mathbb{E}Y \leq_{cx} Y$, then $\mathbb{E}Y = H_g(\mathbb{E}Y) \leq H_g(Y)$.

Because we are interested in risk sharing, cash equivariance is a desirable property because receiving fixed payments (at least within a reasonable range) should not affect attitudes towards risk. The comonotone additivity property represents inability to diversify risks that always move in the same direction.

2.2 Economic objective

Suppose agent i faces a random loss X_i before any risk exchange for $i = 1, 2, \dots, n$. If the collection of agents trades the original allocation \mathbf{X} for the allocation $\mathbf{Y} \in \mathcal{A}(X)$, then the random loss or payout, including transaction costs, of agent i becomes

$$Z_i = Y_i + (a_i + b_i Y_i + c_i \mathbb{E}Y_i) = (1 + b_i)Y_i + a_i + c_i \mathbb{E}Y_i. \quad (4)$$

The additive factor $a_i \geq 0$ is a fixed cost associated with transferring the risk X_i to the coalition of agents (or to a central clearing house); for example, a_i could be the premium that the agent pays to the coalition to eliminate the risk X_i . The multiplicative factor $b_i \geq 0$ represents costs associated with the actual size of the random loss Y_i , for example, investigative costs that could increase proportionally with the size of the loss. The factor $c_i \in \mathbb{R}$ represents costs that reflect the *expected* size of the payout Y_i , for example, hiring claim administrators; c_i is also net of any premium that the agent *receives* in exchange for accepting the risk Y_i , if the premium equals $(1 + \theta)\mathbb{E}Y_i$ as in Arrow [3].

In fact, we might wish to say that $c_i = -(1 + \theta)$, that is, *all* of this part of the cost function arises from premium received. We explore this in examples later in the paper, as well as at the end of this section. Note that our cost function is related to the one considered by Raviv [25]; specifically, Raviv's cost $k(Y_i)$ was a positive, increasing, convex deterministic function of the indemnity Y_i , while allowing for any of those properties to hold weakly. The portion of our cost function given by $a_i + b_i Y_i$ is a special case of such costs; however, the inclusion of the term $c_i \mathbb{E}Y_i$ is not included in Raviv's framework.

Agent i , for $i = 1, 2, \dots, n$, seeks to minimize $H_{g_i}(Z_i)$ for some concave distortion function g_i . Note that minimizing

$$H_{g_i}(Z_i) = H_{g_i}((1 + b_i)Y_i + a_i + c_i \mathbb{E}Y_i) = (1 + b_i)H_{g_i}(Y_i) + a_i + c_i \mathbb{E}Y_i \quad (5)$$

is equivalent to minimizing

$$V_i(Y_i) := (1 + b_i)H_{g_i}(Y_i) + c_i \mathbb{E}Y_i. \quad (6)$$

In light of this recasting of agent i 's goal, a Pareto optimal risk exchange is defined as follows:

Definition 4 $\mathbf{X}^* \in \mathcal{A}(X)$ is called a Pareto optimal risk exchange or allocation if whenever there exists an allocation $\mathbf{Y} \in \mathcal{A}(X)$ such that $V_i(Y_i) \leq V_i(X_i^*)$ for all $i = 1, 2, \dots, n$, then $V_i(Y_i) = V_i(X_i^*)$ for all $i = 1, 2, \dots, n$.

In other words, there is no way to make any agent (strictly) better off without making another agent (strictly) worse off.

We assume that the initial allocation carries finite risk, that is, $H_{g_i}(X_i)$ is finite for $i = 1, 2, \dots, n$. Therefore, there exists at least one allocation \mathbf{Y} , namely \mathbf{X} itself, such that $V_i(Y_i)$ is finite for all $i = 1, 2, \dots, n$.

We end this section by discussing the rationality constraint mentioned in the Introduction. In order that the allocation $\mathbf{Y} \in \mathcal{A}(X)$ be feasible (regardless of whether it is Pareto optimal), it must be true that each agent is at least as well off under \mathbf{Y} as under the original allocation \mathbf{X} . That is, the following inequality must hold for each $i = 1, 2, \dots, n$: $H_{g_i}(X_i) \geq V_i(Y_i)$. We assume that the set of feasible allocations in $\mathcal{A}(X)$ is non-empty.

When first presenting the cost function $a_i + b_i Y_i + c_i \mathbb{E}Y_i$ in connection with Eq. (4), we proposed that one might wish to consider the last term as representing premium received in exchange for accepting the risk Y_i . In that case, write the premium as $-c_i \mathbb{E}Y_i = (1 + \theta) \mathbb{E}Y_i$, so that the rationality constraint becomes

$$(1 + \theta) \mathbb{E}Y_i \geq a_i + (1 + b_i) H_{g_i}(Y_i) - H_{g_i}(X_i). \quad (7)$$

One can interpret the left-hand side of inequality (7) as the minimum premium that agent i is willing to accept for replacing X_i with Y_i . Therefore, the rationality constraint holds in this case if the premium received is at least as great as the risk-adjusted cost, as measured by the right-hand side of (7).

3 Pareto optimal allocations

To describe Pareto optimal allocations, we begin with a basic lemma.

Lemma 1 *If $X^* = (X_1^*, X_2^*, \dots, X_n^*) \in \mathcal{A}(X)$ is Pareto optimal, then so is $(X_1^*, X_2^*, \dots, X_j^* + \beta, \dots, X_k^* - \beta, \dots, X_n^*) \in \mathcal{A}(X)$ for any $\beta \in \mathbb{R}$ and any $j, k = 1, 2, \dots, n$.*

Lemma 1 follows straightforwardly from the cash-equivariance of H_i ; it implies that without loss of generality, we can assume that a Pareto optimal allocation assigns the loss 0 to each of the n agents when the total loss X is 0. If this particular Pareto optimal allocation does not satisfy the rationality constraint in inequality (7), then we can modify the allocation by constants (that sum to zero) so that the rationality constraint is satisfied. (Recall that we assume that the set of feasible allocations is non-empty, so there exist such constants.)

For the present, we skip the degenerate case in which $1 + b_i + c_i = 0$ for all $i = 1, 2, \dots, n$; we consider it more fully for the case of $n = 2$ in Sect. 5. Accordingly, for the remainder of the section we assume that at least one $(1 + b_i + c_i)$ is non-zero. We next use Lemma 1 to characterize the set of Pareto optimal allocations when we view them as points in \mathbb{R}^n via the mapping $F : \mathcal{A}(X) \rightarrow \mathbb{R}^n$ given by $F(\mathbf{Y}) = (V_1(Y_1), V_2(Y_2), \dots, V_n(Y_n))$.

Theorem 1 *We have the following two alternatives:*

1. *If there exist $i, j \in \{1, 2, \dots, n\}$, such that $1 + b_i + c_i \neq 0$ and $(1 + b_i + c_i)(1 + b_j + c_j) \leq 0$, then no Pareto optimal allocation in $\mathcal{A}(X)$ exists.*
2. *Otherwise, the image of the set of Pareto optimal allocations in $\mathcal{A}(X)$ under the mapping F is a hyperplane in \mathbb{R}^n given by*

$$\left\{ \mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n (V_i(X_i^*) - x_i) / (1 + b_i + c_i) = 0 \right\}, \quad (8)$$

in which $\mathbf{X}^* \in \mathcal{A}(X)$ is any Pareto optimal allocation. Furthermore, one obtains such a Pareto optimal allocation \mathbf{X}^* by minimizing

$$\sum_{i=1}^n V_i(Y_i) / |1 + b_i + c_i| \quad (9)$$

over $\mathbf{Y} \in \mathcal{A}(X)$.

Proof First, consider case (a). Without loss of generality, suppose that $1 + b_1 + c_1 < 0$ and $1 + b_2 + c_2 \geq 0$. Consider any $\mathbf{Y} \in \mathcal{A}(X)$. Then, $\mathbf{Z} = (Y_1 + 1, Y_2 - 1, Y_3, \dots, Y_n)$ is a strict improvement on \mathbf{Y} because $V_1(Z_1) = V_1(Y_1) + (1 + b_1 + c_1) < V_1(Y_1)$ and $V_2(Z_2) = V_2(Y_2) - (1 + b_2 + c_2) \leq V_2(Y_2)$. Thus, there cannot exist a Pareto optimal allocation in $\mathcal{A}(X)$.

It follows that the key case is when all $(1 + b_i + c_i)$ are non-zero and of same sign (say positive). We begin by showing that if $\mathbf{X}^* \in \mathcal{A}(X)$ minimizes the expression in (9), then \mathbf{X}^* is Pareto optimal. Suppose that $\mathbf{Y} \in \mathcal{A}(X)$ is such that $V_i(Y_i) \leq V_i(X_i^*)$ for $i = 1, 2, \dots, n$. Then, $\sum_{i=1}^n V_i(Y_i) / |1 + b_i + c_i| \leq \sum_{i=1}^n V_i(X_i^*) / |1 + b_i + c_i|$, from which it follows that $\sum_{i=1}^n V_i(Y_i) / |1 + b_i + c_i| = \sum_{i=1}^n V_i(X_i^*) / |1 + b_i + c_i|$ because \mathbf{X}^* minimizes (9). Therefore, $V_i(Y_i) = V_i(X_i^*)$ for $i = 1, 2, \dots, n$, and \mathbf{X}^* is Pareto optimal.

Suppose $\mathbf{x} \in \mathbb{R}^n$ satisfies the equation of the hyperplane (8) for some Pareto optimal allocation $\mathbf{X}^* \in \mathcal{A}(X)$. Define $\beta_i := (x_i - V_i(X_i^*)) / (1 + b_i + c_i)$ for $i = 1, 2, \dots, n$; then, $\sum_{i=1}^n \beta_i = 0$. Define $\hat{\mathbf{X}}^* := (X_1^* + \beta_1, X_2^* + \beta_2, \dots, X_n^* + \beta_n) \in \mathcal{A}(X)$. By the same argument as in the proof of Lemma 1, one can show that $\hat{\mathbf{X}}^*$ is Pareto optimal. Finally, $F(\hat{\mathbf{X}}^*) = (V_1(X_1^* + \beta_1), V_2(X_2^* + \beta_2), \dots, V_n(X_n^* + \beta_n)) = F(\mathbf{X}^*) + (\beta_1(1 + b_1 + c_1), \beta_2(1 + b_2 + c_2), \dots, \beta_n(1 + b_n + c_n)) = F(\mathbf{X}^*) + (x_1 - V_1(X_1^*), x_2 - V_2(X_2^*), \dots, x_n - V_n(X_n^*)) = \mathbf{x}$. Thus, (any) \mathbf{x} in (8) is an image of a Pareto optimal allocation in $\mathcal{A}(X)$ via the mapping F . As an aside, note that all elements of the hyperplane (8) give the same minimum value in the expression (9).

To complete the proof, we need to show that the hyperplane (8) gives us all the Pareto optimal allocations. Suppose not; suppose that there is a Pareto optimal allocation $\mathbf{Y}^* \in \mathcal{A}(X)$ that is mapped to a point not on the hyperplane (8). Then, by the argument in the above paragraph, any point $\mathbf{y} \in \mathbb{R}^n$ that satisfies $\sum_{i=1}^n (V_i(Y_i^*) - y_i) / (1 + b_i + c_i) = 0$ is the image of a Pareto optimal allocation. Thus, we have two parallel hyperplanes both purporting to be the image (under the mapping F) of Pareto optimal allocations in $\mathcal{A}(X)$. Only one of these hyperplanes will be minimal, a contradiction. Thus, the Pareto optimal allocations in $\mathcal{A}(X)$ correspond to points in the hyperplane (8).

To describe Pareto optimal allocations corresponding to points in the hyperplane (8), it is easier to consider comonotone allocations.

Definition 5 An allocation $\mathbf{Y} \in \mathcal{A}(X)$ is called comonotone if Y_i and X are comonotone for $i = 1, 2, \dots, n$.

Note that if \mathbf{Y} is a comonotone allocation then any two Y_i and Y_j are also pairwise comonotone. Ludkovski and Rüschendorf [22, Proposition 1] show that for V_i preserving the convex order, any integrable non-comonotone allocation $\mathbf{X} \in \mathcal{A}(X)$, $X_i \in L^1(\mathbb{P})$ is dominated by some comonotone \mathbf{X}^* , $V_i(X_i^*) \leq V_i(X_i)$, $i = 1, 2, \dots, n$. This result is essentially based on the comonotone \leq_{cx} -improvement result of Landsberger and Meilijson [21]. Since by assumption $\mathbb{E}X_i \leq H_{g_i}(X_i) < \infty$, $X_i \in L^1$ are integrable and all Pareto optimal allocations are comonotone.

For a comonotone allocation $\mathbf{X} = (f_1(X), f_2(X), \dots, f_n(X))$, Denneberg [14, Proposition 4.5] shows that the functions f_i are continuous on $\text{supp}(X)$ for $i = 1, 2, \dots, n$. Moreover, he shows that f_i may be extended to continuous functions on the entire real line such that $\sum_{i=1}^n f_i(x) = x$ for all $x \in \mathbb{R}$. It follows that we can restrict our attention to finding Pareto optimal allocations in

$$\mathcal{C}(X) \triangleq \left\{ (f_1(X), f_2(X), \dots, f_n(X)) \in \mathcal{A}(X) : \right. \\ \left. f_i \text{ cont., non-decreasing, } \sum_{i=1}^n f_i(x) = x \text{ for } x \in \mathbb{R} \right\}. \quad (10)$$

Comonotonicity implies that an optimal risk allocation necessarily satisfies the mutuality principle, whereby the share of each agent depends only on the total risk X . We now use the above results to explicitly characterize the Pareto optimal allocations.

Theorem 2 Suppose $(1 + b_i + c_i)(1 + b_j + c_j) > 0$ for all $i, j = 1, 2, \dots, n$. Then, $\mathbf{X}^* = (f_1^*(X), f_2^*(X), \dots, f_n^*(X)) \in \mathcal{C}(X)$ is a Pareto optimal allocation if and only if

$$\sum_{i \in \mathcal{I}} (f_i^*)'(t) = 1 \quad \text{for } \mathcal{I} = \operatorname{argmin}_{k=1,2,\dots,n} \frac{(1 + b_k)g_k(S_X(t)) + c_k S_X(t)}{|1 + b_k + c_k|}, \quad (11)$$

and $(f_i^*)'(t) = 0$ otherwise.

Proof From Theorem 1 and Ludkovski and Rüschendorf [22], we know that Pareto optimal allocations correspond to minimizers $\mathbf{X}^* \in \mathcal{C}(X)$ of the expression in (9). As discussed after the proof of Lemma 1 without loss of generality, suppose that the Pareto optimal allocation $\mathbf{X}^* = (f_1^*(X), f_2^*(X), \dots, f_n^*(X))$ is such that $f_i^*(0) = 0$ for $i = 1, 2, \dots, n$.

Suppose $Y = f(X)$ for a continuous, non-decreasing real-valued function f on \mathbb{R}_+ with $f(0) = 0$; then,

$$\begin{aligned} (1 + b)H_g(Y) + c\mathbb{E}Y &= (1 + b) \int_0^1 S_{f(X)}^{-1}(p) dg(p) + c \int_0^1 S_{f(X)}^{-1}(p) dp \\ &= (1 + b) \int_0^1 f[S_X^{-1}(p)] dg(p) + c \int_0^1 f[S_X^{-1}(p)] dp \\ &= (1 + b) \int_0^\infty g[S_X(t)] df(t) + c \int_0^\infty S_X(t) df(t) \\ &= \int_0^\infty [(1 + b)g + c](S_X(t)) df(t), \end{aligned} \quad (12)$$

in which the function $(1 + b)g + c$ is defined on $[0, 1]$ by $[(1 + b)g + c](p) = (1 + b)g(p) + cp$. Thus, minimizing expression (9) is equivalent to minimizing

$$\sum_{i=1}^n \int_0^\infty \frac{[(1 + b_i)g_i + c_i](S_X(t))}{|1 + b_i + c_i|} df_i(t), \quad (13)$$

which is minimized by setting $\sum_{i \in \mathcal{I}} (f_i^*)'(t) = 1$ for $\mathcal{I} = \operatorname{argmin}_{k=1,2,\dots,n} \{(1+b_k)g_k(S_X(t)) + c_k S_X(t)\} / |1 + b_k + c_k|$, and by setting $(f_i^*)'(t) = 0$ otherwise. \square

The above theorem implies that under a Pareto optimal allocation, the risk sharing consists of “tranches” where the risk of each tranche is entirely borne by one agent (ignoring equality in the argmin). As expression (11) shows, the optimal allocation Y_i^* of the i th agent consists of a series of laddered European options on the total risk X . Hence, agent i assumes total responsibility for risk levels where $f_i^*(S_X^{-1}(t)) = 1$, and receives full insurance otherwise. Such risk sharing arrangements are observed in practice in credit derivatives, where the total risk X represents a bond portfolio subject to default risk and the corresponding risk is allocated via credit tranches. These credit tranches can be viewed as optimal insurance contracts for a set of representative investors with varying risk measures.

Remark 1 Theorem 1 and the reduction to comonotone allocations are key steps in our argument since they dramatically simplify the structure of Pareto optimal allocations. Note that the only property used in the proof of Theorem 1 was the cash equivariance of the corresponding risk measures, while the only property used in relation to the comonotonicity improvement of Proposition 1 in Ludkovski and Rüschendorf [22] was consistency of H and \leq_{cx} . On the other hand, Bäuerle and Müller [5] show that any law-invariant convex risk measure, subject to a mild continuity requirement, is consistent with the convex order \leq_{cx} . We, therefore, hypothesize that the conclusion of Theorem 2 will hold for arbitrary law-invariant convex risk measures.

4 Constrained risk sharing

We next consider the related situation for which the risk sharing is subject to regulation. This may arise, for example, in an insurance setting where the risk transfer from buyer to insurer is controlled by a government regulator, or in a financial setting where the party taking on risk is subject to a risk management framework, such as Basel II.

The effect of such regulation is to impose further constraints upon some of the Y_i 's in (9). This of course modifies the resulting Pareto optimal allocations since some of the possible optima become infeasible under the constraint. A similar model was studied by Bernard and Tian [6] under the assumption of a VaR constraint. In our framework where we work with distortion risk measures, we instead postulate constraints of the form

$$H_{h_i}(Y_i) \leq B_i, \quad i = 1, 2, \dots, n,$$

in which H_{h_i} is the regulator's (convex) risk measure on the final risk transfer amount Y_i , and B_i is the corresponding risk threshold for agent i .

We modify the set of allocations $\mathcal{A}(X)$ to account for these constraints. Define the set of *constrained* allocations by

$$\mathcal{A}^c(X) \triangleq \left\{ \mathbf{Y} := (Y_1, Y_2, \dots, Y_n) : X = \sum_{i=1}^n Y_i, \ V_i(Y_i) \text{ finite, } H_{h_i}(Y_i) \leq B_i \right\}.$$

We assume that the set of feasible allocations in $\mathcal{A}^c(X)$ is non-empty. Analogous to Definition 4, $\mathbf{X}^* \in \mathcal{A}^c(X)$ is a *constrained* Pareto optimal allocation if whenever there exists an allocation $\mathbf{Y} \in \mathcal{A}^c(X)$ such that $V_i(Y_i) \leq V_i(X_i^*)$ for all $i = 1, 2, \dots, n$, then $V_i(Y_i) = V_i(X_i^*)$ for all $i = 1, 2, \dots, n$.

The next lemma shows that as in Sect. 3 for unconstrained Pareto optimal allocations, without loss of generality we can restrict our attention to constrained Pareto optimal allocations that are comonotone.

Lemma 2 *If $Y \in \mathcal{A}^c(X)$, then there exists $Y' \in \mathcal{C}(X) \cap \mathcal{A}^c(X)$ that improves it in the partial ordering of Sect. 3.*

Proof Ludkovski and Rüschendorf [22, Proposition 1] show that given an arbitrary allocation $Y \in \mathcal{A}^c(X) \subset \mathcal{A}(X)$, there is a comonotone improvement in the stochastic convex order $Y' \in \mathcal{C}(X)$, that is, $Y'_i \leq_{cx} Y_i$ for $i = 1, 2, \dots, n$. Therefore, the allocation Y' improves Y in the partial ordering of Sect. 3 because V_i preserves the convex order for $i = 1, 2, \dots, n$. Moreover, because H_{h_i} also preserves the convex order, it follows that $H_{h_i}(Y'_i) \leq H_{h_i}(Y_i)$ for $i = 1, 2, \dots, n$ and $Y' \in \mathcal{A}^c(X)$ is still feasible. Thus, $Y' \in \mathcal{C}(X) \cap \mathcal{A}^c(X)$. \square

Note that if the constraining risk measure is not convex, then optimal allocations might not be comonotone. For instance, a VaR constraint at level $\alpha\%$ corresponds to the non-concave distortion function $h(p) = 1_{\{p > \alpha\}}$. Such H_h is not consistent with the \leq_{cx} -order, and therefore Lemma 2 does not apply. Indeed, as explicitly shown by Bernard and Tian [6], the resulting optimal allocation might fail to be comonotone.

By using Lemma 2, we reduce the constrained problem to the same situation as in Theorem 2.

Theorem 3 *The optimal risk allocation for the constrained problem is obtained by finding minimizers of*

$$\sum_{i=1}^n \int_0^\infty \frac{[(1 + b_i)g_i + \lambda_i h_i + c_i](S_X(t))}{|1 + b_i + c_i + \lambda_i|} df_i(t), \quad (14)$$

in which $\lambda_i \geq 0$ is a Lagrange multiplier for the i th constraint, for $i = 1, 2, \dots, n$.

It follows from Theorem 3, that we, again, will obtain a ladder-like optimal contract structure, similar to the tranches in (13). Many cases are possible with respect to which of the λ_i 's are positive (that is, the respective constraint binds) versus zero. In particular, a variety of degeneracies might arise if several constraints bind simultaneously. Instead of considering all these cases for an arbitrary n , in Sects. 5.3 and 6 we focus on a simple example with two agents and one constraint, a setting already taken up in Bernard and Tian [6].

5 The special case of $n = 2$ agents

In this section, we specialize our results to the case for which we have two agents. Suppose an individual (agent 2) is facing an insurable random loss $X_2 = X$ and wants to buy insurance $f(X)$ for all or part of the loss X from an insurer (agent 1 with $X_1 = 0$). In this case, our problem amounts to finding a Pareto optimal allocation $(f^*(X), X - f^*(X))$, in which $f^*(X)$ is the insurer's share of the risk X , and $X - f^*(X)$ is the amount of the risk retained by the individual. Arrow [3] showed that if the premium equals $(1 + \theta)\mathbb{E}f(X)$ with $\theta > 0$ and if the individual seeks to maximize his or her expected utility of wealth, then $f^*(X)$ is deductible coverage (that is, $f^*(X)$ is given functionally by $f^*(x) = (x - d)_+$ for some $d \geq 0$, in which x is a specific value of the random loss X). One could view this risk exchange as Pareto optimal if the insurer's goal were to maximize its expected profits (among other

possible criteria). For more recent work in the area of optimal insurance, see Promislow and Young [24] who extended the work of Arrow to other premium rules and optimality criteria.

We first examine the case for which $1 + b_1 + c_1 = 0 = 1 + b_2 + c_2$. Then, we consider the case for which $(1 + b_1 + c_1)(1 + b_2 + c_2) > 0$.

5.1 $1 + b_1 + c_1 = 0 = 1 + b_2 + c_2$

In this case, we have $c_1 = -(1 + b_1)$ and $c_2 = -(1 + b_2)$. It follows from arguments similar to those in Sect. 3 that the Pareto optimal risk exchanges are given as the minimizers over $\mathbf{Y} \in \mathcal{C}(X)$ of the following expression as λ_1 and λ_2 range over the non-negative reals:

$$\lambda_1(1 + b_1) [H_{g_1}(Y_1) - \mathbb{E}Y_1] + \lambda_2(1 + b_2) [H_{g_2}(Y_2) - \mathbb{E}Y_2], \quad (15)$$

with at least one of λ_1 and λ_2 strictly positive. Without loss of generality, suppose $\lambda_1 > 0$. Also, note that $Y_2 = X - Y_1$; let $f(X)$ denote Y_1 . Then, the Pareto optimal risk exchanges are the minimizers over real-valued, continuous, non-decreasing functions f , with $H_{g_i}(f(X))$ finite for $i = 1, 2$, of the following expression as δ ranges over the non-negative reals:

$$[H_{g_1}(f(X)) - \mathbb{E}f(X)] + \delta [\mathbb{E}f(X) - H_{g_2}(f(X))]. \quad (16)$$

Without loss of generality, we can assume that $f(0) = 0$; otherwise, define \hat{f} by $\hat{f}(x) = f(x) - f(0)$ and note that $H_{g_i}(\hat{f}(X)) - \mathbb{E}\hat{f}(X) = H_{g_i}(f(X) - f(0)) - \mathbb{E}(f(X) - f(0)) = H_{g_i}(f(X)) - \mathbb{E}f(X)$ for $i = 1, 2$.

By following the argument of Theorem 2, the derivative of the optimal function f^* is given by

$$(f^*)'(t) = \begin{cases} 1, & \text{if } g_1(S_X(t)) - S_X(t) < \delta [g_2(S_X(t)) - S_X(t)]; \\ \beta, & \text{if } g_1(S_X(t)) - S_X(t) = \delta [g_2(S_X(t)) - S_X(t)]; \\ 0, & \text{otherwise,} \end{cases} \quad (17)$$

in which $\beta \in [0, 1]$ is arbitrary. If we interpret $g(S_X(t)) - S_X(t)$ as the marginal cost of adding more risk (except for the factor of $1 + b$), then f^* increases if the marginal cost for the insurer is less than the marginal cost for the buyer adjusted by the factor $\delta \geq 0$.

Note that if $0 \leq \delta_1 < \delta_2$, then $f_{\delta_1}^* \leq f_{\delta_2}^*$, in which $f_{\delta_i}^*$ corresponds to the minimizer of (16) for $\delta = \delta_i$, $i = 1, 2$. In other words, as the weight given to buyer's risk preference increases, then the insurer assumes more of the risk.

In the special case for which $\delta = 0$, we seek to minimize $H_{g_1}(f(X)) - \mathbb{E}f(X)$ which is greater than or equal to 0 because g_1 is concave. Thus, $H_{g_1}(f(X)) - \mathbb{E}f(X)$ is minimized by $f^* \equiv r$ for any constant r . If g_1 is strictly concave, then this expression is minimized *uniquely* (up to an additive constant) by $f^* \equiv 0$. If g_1 is not strictly concave, then the minimizer of $H_{g_1}(f(X)) - \mathbb{E}f(X)$ is not necessarily unique. As an example, suppose g_1 is given by AVaR, $g_1(p) = \min(\alpha p, 1)$ for some $\alpha > 1$. Then, for $X \sim \text{Bernoulli}(q)$ with $q > 1/\alpha$, the function f given by $f(0) = r$ and $f(1) = r + 1$ is such that $H_{g_1}(f(X)) - \mathbb{E}f(X) = 0$ for any $r \in \mathbb{R}$. In general, if the distortions are not strictly concave, then it is possible that $g_1(S_X(t)) - S_X(t) = \delta [g_2(S_X(t)) - S_X(t)]$ on a set of positive measure, in which case, f^* will not be unique.

We leave the case for which $1 + b_1 + c_1 = 0 = 1 + b_2 + c_2$ because as the reader will see in the next section, the conclusions that we could draw further from Eq. (17) are similar to the ones we will draw from Eq. (20) below.

$$5.2 (1 + b_1 + c_1)(1 + b_2 + c_2) > 0$$

Let $f(X)$ be the random indemnity that the insurer (agent 1) pays to the buyer (agent 2) in exchange for a premium of $(1 + \theta)\mathbb{E}f(X)$ for some $\theta > 0$, with $f(X)$ and $X - f(X)$ comonotone.

For concreteness, in the notation of this paper, set $a_1 = 0$, $b_1 > 0$, $c_1 = -(1 + \theta)$ and $a_2 = (1 + \theta)\mathbb{E}X$, $b_2 = 0$, $c_2 = -(1 + \theta)$. Thus, the condition $(1 + b_1 + c_1)(1 + b_2 + c_2) > 0$ is equivalent to $b_1 < \theta$.

Under these values for the parameters, the rationality constraint for the insurer in (7) becomes

$$(1 + \theta)\mathbb{E}f(X) \geq (1 + b_1)H_{g_1}(f(X)); \quad (18)$$

that is, the insurer is willing to enter into a contract for which the premium $(1 + \theta)\mathbb{E}f(X)$ is at least as great as the risk-adjusted cost, as measured by $(1 + b_1)H_{g_1}(f(X))$. The rationality constraint for the buyer becomes

$$H_{g_2}(f(X)) \geq (1 + \theta)\mathbb{E}f(X); \quad (19)$$

that is, the risk-adjusted benefit for the buyer from receiving $f(X)$ is greater than the cost $(1 + \theta)\mathbb{E}f(X)$.

It is reasonable to assume that the buyer is “more risk averse” than the insurer in the sense that the buyer’s distortion function is a concave transformation of the insurer’s, or equivalently, $g_2 \geq g_1$. Theorem 2 then implies that the optimal function f^* is given by

$$(f^*)'(t) = \begin{cases} 1, & \text{if } g_1(S_X(t)) - S_X(t) < \frac{\theta - b_1}{\theta(1 + b_1)} [g_2(S_X(t)) - S_X(t)]; \\ \beta, & \text{if } g_1(S_X(t)) - S_X(t) = \frac{\theta - b_1}{\theta(1 + b_1)} [g_2(S_X(t)) - S_X(t)]; \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

in which $\beta \in [0, 1]$ is arbitrary. The function f^* in Eq. (20) is similar in form to the one given in (17), with the arbitrary $\delta \geq 0$ replaced by the fixed $0 < (\theta - b_1)/(\theta(1 + b_1)) < 1$.

From the expression in (20), we can deduce several conclusions. Because $(\theta - b_1)/\theta$ increases as θ increases, the optimal insurance f^* increases as the proportional risk loading θ increases. Also, because $(\theta - b_1)/(1 + b_1)$ decreases as $b_1 < \theta$ increases, the optimal insurance f^* decreases as the insurer’s cost b_1 increases. This makes sense because if the proportional cost of the insurer increases, as measured by b_1 , then the insurer is willing to sell less insurance to the buyer.

If g_2 is replaced by a concave distortion $\hat{g}_2 \geq g_2$, then f^* increases because $g_2(S_X(t)) - S_X(t)$ increases. In other words, as the buyer of insurance becomes more risk averse, then the buyer is willing to purchase more insurance at a given price.

We have the following proposition that tells us when the optimal insurance is deductible insurance. We omit its proof because it is a straightforward application of the expression in (20). Recall from the discussion following Lemma 1 that without loss of generality, we can assume that $f^*(0) = 0$, and we do so in this proposition.

Proposition 1 *If $(g_1(p) - p)/(g_2(p) - p)$ increases for $p \in (0, 1)$, then deductible insurance is optimal, that is,*

$$f^*(x) = (x - d)_+ \quad (21)$$

is optimal with the deductible d given by

$$d = \inf \left\{ t : \frac{g_1(S_X(t)) - S_X(t)}{g_2(S_X(t)) - S_X(t)} \leq \frac{\theta - b_1}{\theta(1 + b_1)} \right\}. \quad (22)$$

If no such d exists, then $f^* \equiv 0$.

Note that if $(g_1(p) - p)/(g_2(p) - p)$ increases for $p \in (0, 1)$, then $(g_1(S_X(t)) - S_X(t))/(g_2(S_X(t)) - S_X(t))$ decreases for $t \geq 0$.

We remark that under the assumptions of Proposition 1, it is straightforward to prove that any allocation $(I(X); X - I(X))$ is dominated in the convex order by the allocation $(|X - d|_+; \min(X, d))$ for d defined by $\mathbb{E}[|X - d|_+] = \mathbb{E}[I(X)]$. Proposition 1 gives the optimal d among all possible deductibles.

Proposition 1 is a generalization of Proposition 3.2 in Jouini et al. [20] who also obtained deductible insurance in the context of law-invariant convex risk measures. In contrast to the proof presented here, their non-constructive method relies on convex duality and only applies in the setting of $L^\infty(\mathbb{P})$.

The following corollaries to Proposition 1 deal with special cases of distortion functions, namely proportional hazards transform and AVaR.

Corollary 1 *If $g_i(p) = p^{c_i}$ for $0 < c_2 < c_1 < 1$, then deductible insurance is optimal.*

Moreover, for the proportional hazards transform, $(g_1(p) - p)/(g_2(p) - p)$ increases from 0 to $(1 - c_1)/(1 - c_2) < 1$. Therefore, if $(1 - c_1)/(1 - c_2) < (\theta - b_1)/(\theta(1 + b_1))$ then full coverage is optimal, which occurs when b_1 is small enough. However, if θ is large, then the rationality constraint in inequality (19) might not hold, so full coverage (even though optimal) might not be feasible. In such cases, we can subtract a fixed amount $a > 0$ from the coverage to make it feasible by the buyer, thereby effectively lowering the benefit and the premium. Finally, note that as c_2 decreases (that is, as the buyer becomes more risk averse), the ratio $(g_1(S_X(t)) - S_X(t))/(g_2(S_X(t)) - S_X(t))$ decreases for a given value of $t \geq 0$, which implies that the deductible decreases (that is, the optimal coverage increases).

Corollary 2 *If $g_i(p) = \min(\alpha_i p, 1)$ for $1 < \alpha_1 < \alpha_2$, then deductible insurance is optimal.*

For the AVaR distortion, $(g_1(p) - p)/(g_2(p) - p)$ increases from $(\alpha_1 - 1)/(\alpha_2 - 1)$ to 1. If $(\alpha_1 - 1)/(\alpha_2 - 1) > (\theta - b_1)/(\theta(1 + b_1))$, then zero coverage is optimal. If $b_1 > 0$ and if $S_X(0) = 1$, then full coverage is never optimal.

We end this section with an example in which we show that deductible coverage as defined in the narrow sense of Eq. (21) is not necessarily optimal.

Example 1 Define the distortions g_1 and g_2 on $[0, 1]$ by

$$g_1(p) = \begin{cases} \frac{9}{8}p, & 0 \leq p \leq \frac{1}{2} \\ \frac{7}{8}p + \frac{1}{8}, & \frac{1}{2} < p \leq 1 \end{cases}; \quad g_2(p) = \begin{cases} \frac{4}{3}p, & 0 \leq p \leq \frac{1}{4} \\ p + \frac{1}{12}, & \frac{1}{4} < p \leq \frac{3}{4} \\ \frac{2}{3}p + \frac{1}{3}, & \frac{3}{4} < p \leq 1. \end{cases}$$

If $X \sim \text{Exp}(1)$, $\theta = 1$, and $b_1 = 1/3$, then one can show that optimal insurance f_1^* satisfies

$$(f_1^*)(t) = \begin{cases} t, & 0 \leq t < \ln \frac{3}{2}, \\ \ln \frac{3}{2}, & \ln \frac{3}{2} \leq t < \ln 3, \\ t + \ln \frac{1}{2}, & t \geq \ln 3. \end{cases}$$

In other words, optimal insurance f_1^* exhibits full coverage up to $\ln(3/2)$ followed by no additional coverage until $\ln 3$, after which the coverage is full at the margin.

5.3 Example with constraints

Regulators of insurance often put constraints on insurance contracts that insurers are allowed to provide in the market. To illustrate the effect of constraints on the form of the indemnity contract f , we include a simple example. We follow the model for two agents with $b_1 > 0$, $b_2 = 0$, and $c_1 = c_2 = -(1 + \theta)$. Let

$$\begin{cases} g_1(p) = \min(\alpha_1 p, 1) & h_1(p) = \min(\beta p, 1) \\ g_2(p) = \min(\alpha_2 p, 1) \end{cases}.$$

Agent 1 is the insurer with the AVaR distortion function g_1 that faces a regulator constraint based on the H_{h_1} risk measure; agent 2 is the buyer with the AVaR distortion function g_2 .

Suppose that $\alpha_2 > \beta > \alpha_1 > 1$; that is, the buyer is the most risk averse with the insurer being the least risk averse and the regulator somewhere in between. The relevant terms in the sum (14) are given by

$$\begin{cases} Q_1(p) = [(1 + b_1) \min(\alpha_1 p, 1) - (1 + \theta)p + \lambda \min(\beta p, 1)]/|b_1 + \lambda - \theta|, \\ Q_2(p) = [\min(\alpha_2 p, 1) - (1 + \theta)p]/\theta. \end{cases} \quad (23)$$

By Theorem 3, for a given Lagrange multiplier $\lambda \geq 0$, the optimal contract satisfies $(f^\lambda)'(S_X(t)) = 1$ if $Q_1(p) < Q_2(p)$ and $(f^\lambda)'(S_X(t)) = 0$ otherwise.

In the following, we assume that $\theta > \lambda + b_1$, so the transaction costs are *large*. The risk functions Q_1 and Q_2 are illustrated in Fig. 1. We note that for large $p \sim 1$, $Q_1(p) \geq Q_2(p)$ and moreover, the two piecewise linear functions cross at most once on $(0, 1)$. More precisely, if $\alpha_2 > (1 + \theta) + \frac{[(1+b_1)\alpha_1 - (1+\theta) + \lambda\beta]\theta}{\theta - b_1 - \lambda}$, then for small $p \sim 0$, $Q_1(p) < Q_2(p)$, and Q_1 and Q_2 have exactly one crossing point $0 < p^* < 1$. Thus, the optimal contract in that case is deductible insurance $f^\lambda(x) = (x - d)_+$, as the insurer covers large risks (small p) and the buyer takes on small risks. If α_2 is smaller than the above threshold, then $Q_1(p) > Q_2(p)$ for all $p \in (0, 1)$, and it is optimal to have zero insurance $f^\lambda \equiv 0$ (note that zero insurance implies $\lambda = 0$ as the constraint is necessarily non-binding).

The two (finite) possibilities for the deductible level d (with $S_X(d)$ corresponding to the unique crossing point of Q_1 and Q_2) are illustrated in Fig. 1. The top panel of Fig. 1 shows Case (a), whereby

$$S_X(d) = p_2^* = \frac{\lambda(1 + \theta) - \theta + b_1}{(1 + \theta)(b_1 + \lambda) - (1 + b_1)\alpha_1\theta}. \quad (24)$$

The necessary and sufficient condition for Case (a) to occur is $1/\beta < p_2^* < 1/\alpha_1$, which is equivalent to

$$-b_1 < \lambda < \min\left(\theta - b_1, \frac{(\theta - b_1)\beta + (1 + \theta)b_1 - (1 + b_1)\alpha_1\theta}{(1 + \theta)(\beta - 1)}\right).$$

It is possible that the upper bound is negative which implies that case (a) cannot occur as λ is non-negative by construction.

Otherwise, we are in Case (b) shown on the bottom panel of Fig. 1, where

$$S_X(d) = p_1^* = \frac{\theta - (b_1 + \lambda)}{\theta(1 + b_1)\alpha_1 - b_1(1 + \theta) + \lambda[\theta\beta - (1 + \theta)]}. \quad (25)$$

Case (b) requires that $1/\alpha_2 < p_1^* < 1/\beta$, or

$$\frac{b_1(1 + \theta) - \theta(1 + b_1)\alpha_1 + (\theta - b_1)\beta}{(\beta - 1)(1 + \theta)} < \lambda < \frac{b_1(1 + \theta) - \theta(1 + b_1)\alpha_1 + (\theta - b_1)\alpha_2}{(\alpha_2 - 1) + \theta(\beta - 1)}.$$

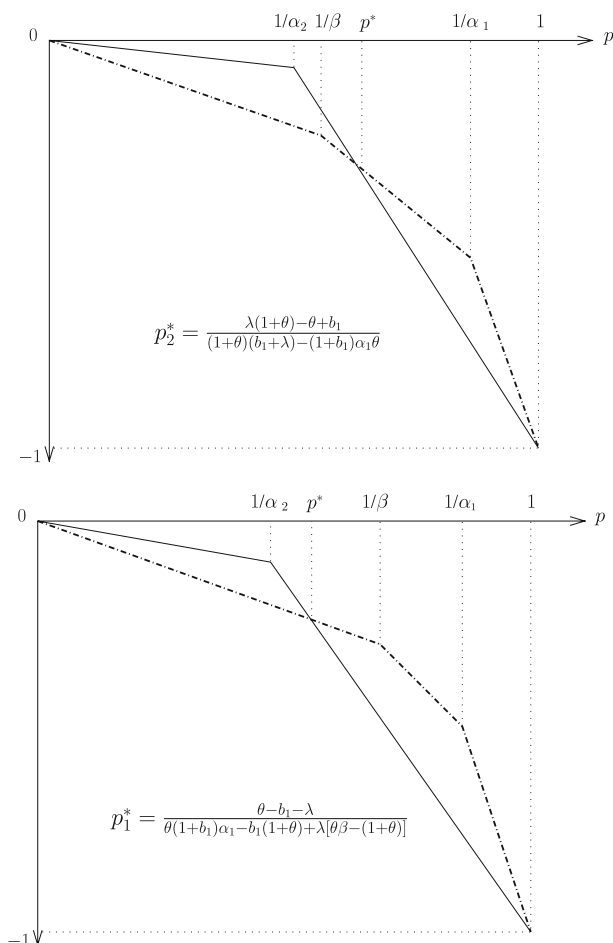


Fig. 1 Risk functions for the example in Sect. 5.3. The *dashed line* represents $Q_1(p) = [(1+b_1)\min(\alpha_1 p, 1) - (1+\theta)p + \lambda\min(\beta p, 1)]/|b_1 + \lambda - \theta|$, and the *solid line* is $Q_2(p) = [\min(\alpha_2 p, 1) - (1+\theta)p]/\theta$. In this example, $\theta > \lambda + b_1$, so we have $Q_1(1) = Q_2(1) = -1$. Note that both functions are piecewise linear. The *crossing points* correspond to the *tranche levels* of optimal contracts. Case (a) is on the *top*, Case (b) is on the *bottom*

Finally, note that if $\beta > \alpha_2 > \alpha_1 > 1$ then it is possible that Q_1 and Q_2 cross twice in the interior of $(0, 1)$, so that the optimal contract may be a capped deductible.

6 Minimizing the risk of the buyer subject to a constraint

To further explore the implications of constrained risk sharing, we consider a slightly different example in which the buyer is the only minimizing agent. This is the usual insurance setting whereby the insurer offers a menu of contracts and the buyer selects the one most suited to her needs. Thus, the optimization is from the buyer's point of view; the insurer's risk preferences enter the problem through the insurance price. This example further illustrates the intricate connection between distortion functions and optimal contract shape.

Assume that the buyer's risk-adjusted loss after obtaining insurance is $(1 + b)H_g(X - f(X)) + (1 + \theta)\mathbb{E}f(X)$, in which the first term represents the residual risk and the second term represents the insurance premium. The insurer himself is constrained by regulators to $H_h(f(X)) \leq B$, so that only a limited amount of risk may be transferred. We ignore the desires of the insurer and focus on minimizing the risk-adjusted loss of the buyer subject to this constraint. Then, we seek to find a non-decreasing f^* that minimizes

$$(1 + b)H_g(X - f(X)) + (1 + \theta)\mathbb{E}f(X), \quad (26)$$

subject to the regulatory constraint

$$H_h(f(X)) \leq B, \quad (27)$$

for some $B > 0$. The following proposition is a direct counterpart of Theorem 3.

Theorem 4 *An insurance contract f^* that minimizes (26) subject to (27) is determined by*

$$(f^*)'(t) = \begin{cases} 1, & \text{if } (1 + b)g(S_X(t)) > (1 + \theta)S_X(t) + \lambda h(S_X(t)), \\ 0, & \text{if } (1 + b)g(S_X(t)) \leq (1 + \theta)S_X(t) + \lambda h(S_X(t)). \end{cases} \quad (28)$$

Furthermore, either $\lambda = 0$ or $\lambda > 0$, with the latter implying that (27) holds with equality, from which we can determine λ .

Proof Fix $\lambda \geq 0$. Proceeding as in (12), we have

$$\begin{aligned} & (1 + b)H_g(X - f^\lambda(X)) + (1 + \theta)\mathbb{E}Y + \lambda(H_h(f^\lambda(X)) - B) \\ &= \int_0^\infty [-(1 + b)g + (1 + \theta) + \lambda h](S_X(t)) df^\lambda(t) + \text{Const.} \end{aligned}$$

Thus, to minimize (26) we should set $(f^\lambda)'(t) = 0$ when the integrand is positive, and $f'(t) = 1$ when the integrand is negative, which is equivalent to (28). To find λ , we solve for $\lambda \int_0^\infty h(S_X(t)) df^\lambda(t) = B$. \square

To be concrete, take $g(p) = \min(\alpha p, 1)$ and $h(p) = \min(\beta p, 1)$, in which $\alpha > \beta > 1$ so that the buyer is more risk averse than the regulator. Also, suppose the loss X is exponentially distributed with mean equal to $1/\mu$. Then, for a given Lagrange multiplier $\lambda \geq 0$, we find f^λ to minimize

$$\begin{aligned} & \int_0^{\frac{1}{\mu} \ln \beta} [-(1 + b)e^{\mu t} + \lambda e^{\mu t} + (1 + \theta)] e^{-\mu t} df(t) \\ &+ \int_{\frac{1}{\mu} \ln \beta}^{\frac{1}{\mu} \ln \alpha} [-(1 + b)e^{\mu t} + \lambda \beta + (1 + \theta)] e^{-\mu t} df(t) \\ &+ \int_{\frac{1}{\mu} \ln \alpha}^\infty [-(1 + b)\alpha + \lambda \beta + (1 + \theta)] e^{-\mu t} df(t). \end{aligned} \quad (29)$$

Depending on the signs of the three integrands in (29) we obtain five possible cases for the relationship between b , θ and λ . Further sub-cases are then obtained depending on the

Table 1 Classification of Pareto optimal allocations of example in Sect. 6

$\theta > (1+b)\alpha - 1$			
$B > 0$	Case 5	$d = +\infty$	$\lambda = 0$
$\theta = (1+b)\alpha - 1$			
$B > 0$	Case 4a	non-unique optimum	$\lambda = 0$
$(1+b)\beta - 1 \leq \theta < (1+b)\alpha - 1$			
$\mu B \leq \beta/\alpha$	Case 4b	non-unique optimum	$\lambda = ((1+b)\alpha - (1+\theta))/\beta$
$\beta/\alpha < \mu B < \frac{(1+b)\beta}{1+\theta}$	Case 3b	$d = (1/\mu) \ln \left(\frac{\beta}{\mu B} \right)$	$\lambda = \frac{1+b}{\mu B} - \frac{1+\theta}{\beta} > 0$
$\mu B \geq \frac{(1+b)\beta}{1+\theta}$	Case 3a	$d = (1/\mu) \ln \left(\frac{1+\theta}{1+b} \right)$	$\lambda = 0$
$b < \theta < (1+b)\beta - 1$			
$\mu B \leq \beta/\alpha$	Case 4b	non-unique optimum	$\lambda = ((1+b)\alpha - (1+\theta))/\beta$
$\beta/\alpha < \mu B < 1$	Case 3b	$d = (1/\mu) \ln \left(\frac{\beta}{\mu B} \right)$	$\lambda = \frac{1+b}{\mu B} - \frac{1+\theta}{\beta} > 0$
$1 \leq \mu B < 1 + \ln \left(\frac{\beta(1+b)}{1+\theta} \right)$	Case 2b2	$d = -B + \frac{1+\ln \beta}{\mu}$	$\lambda = (1+b) - \frac{1+\theta}{\beta} e^{\mu B - 1}$
$\mu B \geq 1 + \ln \left(\frac{\beta(1+b)}{1+\theta} \right)$	Case 2a	$d = 1/\mu \ln \left(\frac{1+\theta}{1+b} \right)$	$\lambda = 0$
$\theta \leq b$			
$\mu B \leq \beta/\alpha$	Case 4	non-unique optimum	$\lambda = ((1+b)\alpha - (1+\theta))/\beta$
$\beta/\alpha < \mu B < 1$	Case 3b	$d = (1/\mu) \ln \left(\frac{\beta}{\mu B} \right)$	$\lambda = \frac{1+b}{\mu B} - \frac{1+\theta}{\beta} > 0$
$1 \leq \mu B < 1 + \ln \beta$	Case 2b1	$d = -B + \frac{1+\ln \beta}{\mu}$	$\lambda = (1+b) - \frac{1+\theta}{\beta} e^{\mu B - 1}$
$\mu B \geq 1 + \ln \beta$	Case 1	$d = 0$	$\lambda = 0$

value of regulator's B which controls whether the constraint is binding $\lambda > 0$ or not. The full results are summarized in Table 1, see also Ludkovski and Young [23] for further details.

An interesting sub-case occurs when $-(1+b)\alpha + \lambda\beta + (1+\theta) = 0$ and $\theta < (1+b)\alpha - 1$. In this case $\lambda = ((1+b)\alpha - (1+\theta))/\beta > 0$, from which it follows that $\lambda + \theta - b > 0$ and $-(1+b)\beta + \lambda\beta + (1+\theta) > 0$. Thus, the first two integrals in (29) are positive, which implies that $(f^\lambda)'(t) = 0$ for $t \leq (\ln \alpha)/\mu$. Moreover, the third integral is identically zero, so we have infinitely many possible solutions f^λ , even though the constraint binds. This degeneracy arises due to the piecewise linear form of the AVaR distortions we selected. Thus, f^λ is given by $(f^\lambda)'(t) = 0$ for $t < (\ln \alpha)/\mu$ and arbitrary $(f^\lambda)'(t) \in [0, 1]$ for $t \geq (\ln \alpha)/\mu$ such that $H_h(f^\lambda(X)) = B$. We give two examples to illustrate possible indemnity functions f^λ :

- i: Let $f^\lambda(x) = r(x - (\ln \alpha)/\mu)_+$ with proportional coverage $r = \mu B \alpha / \beta$. Note that $r \in [0, 1]$ if and only if $\mu B \leq \beta/\alpha$.
- ii: Let $f^\lambda(x) = \min(r'(x - d')_+, m - d')$ with d' and r' given such that $d' \geq (\ln \alpha)/\mu$ and $\mu B e^{\mu d'}/\beta < r' \leq 1$, from which it follows that $m = (1/\mu) \ln \left(\frac{\beta r' e^{\mu d'}}{\beta r' - \mu B e^{\mu d'}} \right) > d'$.

7 Summary and conclusions

In this paper, we proved that (Pareto) optimal risk sharing contracts take the form of deductible insurance in the setting of agents endowed with distortion risk measures and linear transaction/premium costs. Such results continue to hold under third-party constraints. This conforms to real-life insurance contracts both in a two-agent case (for example, casualty reinsurance) and in a multi-agent setting (credit derivatives based on tranches).

It would be interesting to extend our results to a more general setting, in particular indifference measures based on Rank Dependent Expected Utility (also known as Maximin Expected Utility). A tractable example is the exponential-distortion risk measure, see Tsanakas and Desli [29]:

$$H(X) = \frac{1}{\gamma} \ln \left\{ \int_{-\infty}^0 (g[S_{e^{\gamma}}(t)] - 1) dt + \int_0^{\infty} g[S_{e^{\gamma}}(t)] dt \right\}. \quad (30)$$

The preferences induced by H can be seen in the context of robust utility, where the parameter γ is interpreted as the risk aversion coefficient, while the distortion function g corresponds to ambiguity-aversion. One can show that H is a law-invariant, convex risk measure. However, compared to our model (2), H is no longer coherent or comonotone additive. Nevertheless, by Remark 1 our analysis up to Theorem 2 still applies. However, because the non-linear log-transformation in (30) is global, the structure of Theorem 2 does not hold because we can no longer perform t -by- t optimization for the optimal risk allocation f . Related analysis of RDEU risk-sharing was carried out by Carlier and Dana [9, 10] using differential methods; these authors obtained a mixture of co-insurance and deductible ladders.

From a general viewpoint, our work confirms previous results of Jouini et al. [20] (and originally Arrow [3]) on optimality of deductible insurance. Conversely, it contrasts with possibility of proportional risk sharing obtained in Barrieu and El Karoui [4] (and originally Borch [7]). The key step in our method relies on comonotonicity of Pareto optimal allocations due to the consistency of preferences with the stochastic convex order \leq_{cx} . Thus, we raise the conjecture that in the setting of law-invariant convex risk measures, optimal risk sharing always leads to insurance that incorporates a ladder of deductibles (both in unconstrained and constrained settings).

Acknowledgments We thank Carole Bernard and Damir Filipovic for useful discussions. We are also grateful to the editor and the anonymous referees for their many suggestions that have greatly improved our presentation.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Aase, K.K.: Perspectives of risk sharing. *Scand. Actuar. J.* **2**, 73–128 (2002)
2. Acciaio, B.: Optimal risk sharing with non-monotone monetary functions. *Finance Stoch.* **11**(2), 267–289 (2007)
3. Arrow, K.J.: Uncertainty and the welfare of medical care. *Am. Econ. Rev.* **53**, 941–973 (1963)
4. Barrieu, P., El Karoui, N.: Inf-convolution of risk measures and optimal risk transfer. *Finance Stoch.* **9**(2), 269–298 (2005)
5. Bäuerle, N., Müller, A.: Stochastic orders and risk measures: consistency and bounds. *Insurance: Math. Econ.* **38**(1), 132–148 (2006)
6. Bernard, C., Tian, W.: Insurance market effects of risk management metrics. *Geneva Risk Insur. Rev.* (forthcoming) (2009)
7. Borch, K.: Equilibrium in a reinsurance market. *Econometrica* **30**(3), 424–444 (1962)
8. Burgert, C., Rüschendorf, L.: Allocation of risks and equilibrium in markets with finitely many traders. *Insurance: Math. Econ.* **42**(1), 177–188 (2008)
9. Carlier, G., Dana, R.-A.: Are generalized call-spreads efficient? *J. Math. Econ.* **43**(5), 581–596 (2007)
10. Carlier, G., Dana, R.-A.: Two-persons efficient risk-sharing and equilibria for concave law-invariant utilities. *Econ. Theory* **36**(2), 189–223 (2008)

11. Chateauneuf, A., Dana, R.-A., Tallon, J.-M.: Optimal risk-sharing rules and equilibria with Choquet-expected-utility. *J. Math. Econ.* **34**(2), 191–214 (2000)
12. Dana, R.-A., Meilijson, I.: Modelling agents' preferences in complete markets by second order stochastic dominance, working paper, Cahier du Ceremade 0238 (2003)
13. Denuit, M., Dhaene, J., Goovaerts, M., Kaas, R., Laeven, R.J.A.: Risk measurement with equivalent utility principles. Working Paper, K.U. Leuven (2006)
14. Denneberg, D.: Non-Additive Measure and Integral. Kluwer, Dordrecht (1994)
15. Filipovic, D., Kupper, M.: Optimal capital and risk transfers for group diversification. *Math. Finance* **18**(1), 55–76 (2008)
16. Filipovic, D., Kupper, M.: Equilibrium prices for monetary utility functions. *Int. J. Theor. Appl. Finance* **11**, 325–343 (2008)
17. Föllmer, H., Schied, A.: Stochastic Finance. An Introduction in Discrete Time, 2nd edn. Walter de Gruyter, Amsterdam (2004)
18. Gerber, H.U.: Pareto-optimal risk exchanges and related decision problems. *ASTIN Bull.* **10**(1), 25–33 (1978)
19. Jouini, E., Schachermayer, W., Touzi, N.: Law invariant risk measures have the Fatou property. *Adv. Math. Econ.* **9**, 49–72 (2006)
20. Jouini, E., Schachermayer, W., Touzi, N.: Optimal risk sharing for law invariant monetary utility functions. *Math. Finance* **18**(2), 269–292 (2008)
21. Landsberger, M., Meilijson, I.: Co-monotone allocations, Bickel–Lehmann dispersion and the Arrow–Pratt measure of risk aversion. *Ann. Oper. Res.* **52**, 97–106 (1994)
22. Ludkovski, M., Rüschendorf, L.: On comonotonicity of Pareto optimal allocations. *Stat. Probab. Lett.* **78**(10), 1181–1188 (2008)
23. Ludkovski, M., Young, V.R.: Optimal risk sharing under distorted probabilities. Preprint (2008). [arxiv.org/0809.3778](https://arxiv.org/abs/0809.3778)
24. Promislow, S.D., Young, V.R.: Unifying framework for optimal insurance. *Insurance: Math. Econ.* **36**(3), 347–364 (2005)
25. Raviv, A.: The design of an optimal insurance policy. *Am. Econ. Rev.* **69**(1), 84–96 (1979)
26. Rothschild, M., Stiglitz, J.E.: Increasing risk, I: a definition. *J. Econ. Theory* **2**, 225–243 (1970)
27. Rothschild, M., Stiglitz, J.E.: Increasing risk, II: its economic consequences. *J. Econ. Theory* **3**, 66–84 (1971)
28. Rothschild, M., Stiglitz, J.E.: Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Q. J. Econ.* **90**, 629–649 (1976)
29. Tsanakas, A., Desli, E.: Risk measures and theories of choice. *Br. Actuar. J.* **9**(4), 959–991 (2003)
30. Wang, S.S., Young, V.R.: Ordering risks: expected utility theory versus Yaari's dual theory of risk. *Insurance: Math. Econ.* **22**(2), 145–161 (1998)
31. Wang, S.S., Young, V.R., Panjer, H.H.: Axiomatic characterization of insurance prices. *Insurance: Math. Econ.* **21**, 173–183 (1997)
32. Wilson, R.: The theory of syndicates. *Econometrica* **36**(1), 119–132 (1968)
33. Yaari, M.E.: The dual theory of choice under risk. *Econometrica* **55**, 95–115 (1987)
34. Young, V.R., Browne, M.J.: Equilibrium in competitive insurance markets under adverse selection and Yaari's dual theory of risk. *Geneva Pap. Risk Insurance Theory* **25**, 141–157 (2000)